

BIOINFORMATICS ANALYSIS OF ZMBBI GENE FAMILY OF MAIZE (*ZEA MAYS* L.)

LONG JIANG, JINGZHE ZHANG, XIAOMING YU*, JIAXIN LIU,
HAIYAN YU AND JIANMING LI

*College of Agronomy, Jilin Agricultural Science and Technology University,
Jilin, Jilin 132101, Jilin Province, P. R. China*

Keywords: Maize, ZmBBI, E3 linkase, Bioinformatics analysis

Abstract

Seven AtBBI and 17 ZmBBI genes were identified, and their phylogenetic tree members, the conservative structure along with the physical and chemical properties of ZmBBI protein were analyzed by bioinformatics method. The number of amino acids, molecular length, theoretical isoelectric point value, fat coefficient value and hydrophobic average coefficient value of ZmBBI family were also analyzed. In the subcellular location, all ZmBBI genes were located in the nucleus. The tissue-specific expression of ZmBBI gene family showed that the expression of different genes in different tissues and growth stages were different. Zm00001d052154 and Zm00001d037167 were expressed in the most tissues and growth stages. The expression of Zm00001d04831 was not concentrated in some tissues. Zm00001d021225 was highly expressed in corn silk and cortex. Zm00001d021314 was well expressed not only in the whole process of corn growth and development, but also in adversity. In addition, the specific analysis of ZmBBI gene in various tissues and the abiotic stress response expression analysis under the condition of multi salt and high permeability showed that Zm00001d039623 and Zm00001d009571 were both highly expressed in mature pollen, and Zm00001d039623 was not expressed under salt stress and high permeability stress. The research results are for further study on the function of ZmBBI gene family.

Introduction

Maize (*Zea mays* L.) an important food, feed, energy and commercial crop in the world plays an important role in agricultural production and national economy (Zhao *et al.* 2018, Jiang *et al.* 2020a). In recent years, the consumption of maize as food is increasing every year, and the demand for corn as food and industrial use is also increasing (Jiang *et al.* 2020b).

Blast and Bth-induced 1 OsBBI1 gene encodes the activity of RING-finger protein and E3 binding enzyme, which determines the tolerance of rice to virus (Li 2010). Biochemical experiments showed that OsBBI protein had E3 binding enzyme activity after being extracted in laboratory (Li 2010). Gene analysis showed that the loss of TZm17 function in the ZmBBI implanted mutant increased the sensitivity of rice to the virus. It indicated that ZmBBI regulated the wide tolerance of rice to rice blast virus. The cell wall of the over expressed plants was thicker than that of the mutant plants. ZmBBI plays an important role in the regulation of gene expression. At present, there are no such reports and articles, so it is necessary to analyze the family and further understand its role.

Based on the translated protein data of maize genome, this study analyzed the sequence characteristics, gene expression, chromosome position information and phylogeny of ZmBBI gene family by using bioinformatics related methods, which provided a good direction for the follow-up functional analysis of ZmBBI in maize.

*Author for correspondence: <yuxm446@nenu.edu.cn>.

Materials and Methods

From the online database PlantTFDB (<http://plantfdb.cbi.pku.edu.cn/>), the protein sequences of ZmBBI gene family and AtBBI gene family are used as the premise of other related results in this study (Liu *et al.*, 2015, Jin *et al.* 2017, Jiang *et al.* 2020). From online database PlantTFDB (<http://plantfdb.cbi.pku.edu.cn/>), the amino acid sequence of 1 OsBBI1 (LOC_ Os06g03580) in rice, 7 Arabidopsis genes (At3g63530 (ctl19), At5g52140 (ctl17), At3g19910 (ctl18), At3g47180 (ctl16), At2g15530 (ctl03), At1g53190 (ctl01), At4g34040 (ctl04) and 17 ZmBBIs in maize (Zm00001d021225, Zm00001d002624, Zm00001d045517, Zm00001d035989, Zm00001d048301, Zm00001d027868, Zm00001d053212, Zm00001d021314, Zm00001d039769, Zm00001d018190, Zm00001d008649, Zm00001d052154, Zm00001d037167, Zm00001d048796, Zm00001d039623, Zm00001d018328, and Zm00001d009571) were obtained.

Seventeen ZmBBI proteins, 7 AtBBI proteins and 1 OsBBI protein were aligned by using ClustalW (Jeanmougin *et al.* 1998), and the phylogenetic tree was constructed by Software MEGA 5.05, which was repeated 1000 times, and the remaining parameters remained unchanged compared to “Default” setting of the tool (Zhao *et al.* 2015).

Using the NCBI’s online program Conserved Domain Database (<https://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi>), the conserve domain of each ZmBBI in maize was analyzed (Lu *et al.* 2020). According to the available genomic location information, the location of each gene in 17 ZmBBI family genes on 10 different maize chromosomes was analyzed and found. Based on MaizeGDB database (<https://www.maizegdb.org/>) BLSAT was used to find the location of each ZmBBI gene on each chromosome. Mainly using online analysis software ProtParam provided by the bioinformatics resource portal “ExPASy” of the SIB Swiss Institute of Bioinformatics (<http://web.expasy.org/protparam/>), the amino acid number, molecular weight length, theoretical isoelectric point number, aliphatic amino acid number and protein hydrophobicity of 17 maize ZmBBI family genes were analyzed (Artimo *et al.* 2012, Xie *et al.* 2014).

Using Plant-mPLOC - an online website for plant subcellular localization (<http://www.csbio.sjtu.edu.cn/bioinf/Plant-multi/>), the location of 17 ZmBBI proteins in cells was determined (Chou *et al.* 2010, Zhu *et al.* 2012).

Using qTeller for MaizeGDB (https://qteller.maizegdb.org/genes_by_name_B73v4.php), the expression level and change pattern of 17 ZmBBI genes in different maize tissue regions and plant development stages were analyzed. The mapping of gene expression was made by using the mapping headset function in TBtools (Jiang *et al.* 2020a, 2020b).

Results and Discussion

Based on the analysis the database PlantTFDB, one OsBBI protein, seven AtBBI proteins and 17 ZmBBI proteins were finally obtained. The 17 ZmBBI proteins of maize were as follows: Zm00001d021225 (GRMZM2G073228), Zm00001d002624 (GRMZM2G081060), Zm00001d045517 (GRMZM2G071602), Zm00001d035989 (GRMZM2G099278), Zm00001d048301 (GRMZM2G141084), Zm00001d027868 (GRMZM2G021498), Zm00001d053212 (GRMZM2G477205), Zm00001d021314 (GRMZM2G085948), Zm00001d039769 (GRMZM2G165044), Zm00001d018190 (GRMZM2G021480), Zm00001d008649 (GRMZM2G305901), Zm00001d052154 (GRMZM2G473016), Zm00001d037167 (GRMZM2G142816), Zm00001d048796 (GRMZM2G468260), Zm00001d039623 (GRMZM2G004799), Zm00001d018328 (GRMZM2G089466), and Zm00001d009571 (GRMZM2G124701), respectively (Table 1).

ZmBBI proteins were compared with each other by aligning using ClustalW. Phylogenetic study revealed that 25 ZmBBI proteins in maize, rice and Arabidopsis are clustered into two parts

and named as Cluster-1 and Cluster-2 (Fig. 1). It was found that there were 7 ZmBBI proteins and 3 AtBBI proteins in Cluster-1, while in Cluster-2, there were 1 OsBBI protein, 10 ZmBBI proteins and 4 AtBBI proteins, indicating that BBI proteins expanded and differentiated before the divergence of rice and Arabidopsis.

Table 1. Length of ZmBBI gene.

Gene	Locus	Gene location
Zm00001d021225	GRMZM2G073228	Chr7:146550796-146555239
Zm00001d002624	GRMZM2G081060	Chr2:17074071-17087531
Zm00001d045517	GRMZM2G071602	Chr9:25249729-25252528
Zm00001d035989	GRMZM2G099278	Chr6:64968486-64975065
Zm00001d048301	GRMZM2G141084	Chr9:154161995-154167015
Zm00001d027868	GRMZM2G021498	Chr1:15763100-15768449
Zm00001d053212	GRMZM2G477205	Chr4:220417152-220419734
Zm00001d021314	GRMZM2G085948	Chr7:148630663-148635120
Zm00001d039769	GRMZM2G165044	Chr3:14333996-14339635
Zm00001d018190	GRMZM2G021480	Chr5:216100153-216101154
Zm00001d008649	GRMZM2G305901	Chr8:15902706-15910406
Zm00001d052154	GRMZM2G473016	Chr4:181429730-181431252
Zm00001d037167	GRMZM2G142816	Chr6:114642767-114643789
Zm00001d048796	GRMZM2G468260	Chr4:5651769-5655950
Zm00001d039623	GRMZM2G004799	Chr3:9345334-9348991
Zm00001d018328	GRMZM2G089466	Chr5:218514190-218518512
Zm00001d009571	GRMZM2G124701	Chr8:71287101-71293262

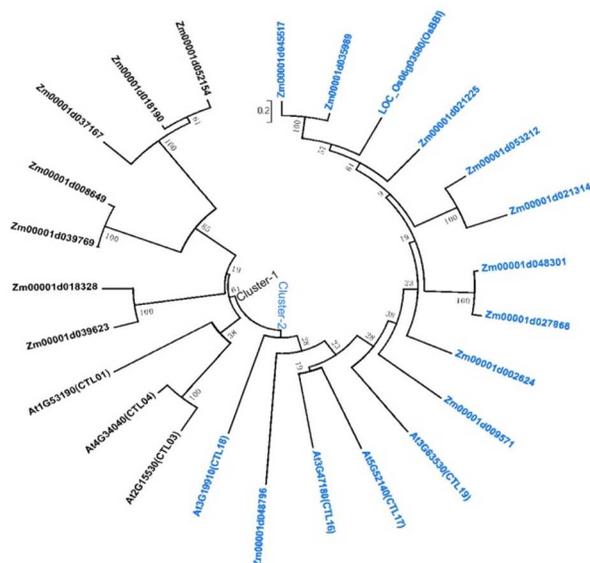


Fig. 1. Phylogenetic tree of OsBBI, ZmBBI and AtBBI genes. Amino acid sequences of one OsBBI protein, seven AtBBI proteins and 17 ZmBBI proteins are obtained from database Plant TFDB, and by aligning using ClustalW. The color of genes' names in Cluster-1 are black and they are blue in the Cluster-2.

Ring finger domain and U-box domain are special types of zinc fingers, which have many residues and combine two zinc atoms. Its molecule is defined by the "cross support" mode (Zheng 2013). The phase mode is composed of 8 amino acid residues to synthesize a zinc atom, usually Cys or his, which is arranged by the characteristic distance. The structure of U-box is similar to that of 3 β plates and 1 α helix, which is stable by salt bridge. Many of them are ubiquitin protein ligase (E3s), as a scaffold binding ubiquitin carrier protein substrate protein, which enables ubiquitin to be effectively transferred from E2 to substrate (Jiang 2010). The analysis of the relationship between the conserved domains of several ZmBBI protein sequences shows that Zm00001d052154, Zm00001d018190 and Zm00001d037167 of Cluster-1 contain a Zinc structure and a Ring_ubox structure, while Zm00001d008649, Zm00001d039769, Zm00001d018328 and Zm00001d039623 contain a Ring_Ubox structure. In Cluster-2, Zm00001d021225, Zm00001d045517, Zm00001d04831 and Zm00001d027868 contain a COG5219 structure, and Zm00001d048796, Zm00001d053122, Zm00001d009571 and Zm00001d002624 all contain a Ring_Ubox structure.

Table 2. Analysis of physical and chemical properties and subcellular location of ZmBBI gene family proteins.

Gene locus	Amino acid number	Molecular weight	Theoretical ISO electric point	Fat co-efficient	Av. co-efficient of hydrophobicity	Predicted location
Zm00001d021225	175	19737.1	5.32	71.94	-0.568	Nucleus
Zm00001d002624	336	37905.7	6.08	77.50	-0.459	"
Zm00001d045517	268	29215.66	4.39	61.57	-0.691	"
Zm00001d035989	647	70755.87	5.18	86.77	-0.252	"
Zm00001d048301	241	27211.33	6.64	59.05	-0.704	"
Zm00001d027868	270	30815.62	5.60	63.89	-0.459	"
Zm00001d053212	287	30776.76	3.75	61.99	-0.728	"
Zm00001d021314	322	35301.75	4.08	62.20	-0.837	"
Zm00001d039769	551	59134.64	8.61	72.11	-0.331	"
Zm00001d018190	333	36531.3	8.78	58.95	-0.517	"
Zm00001d008649	527	57094.11	8.26	69.47	-0.468	"
Zm00001d052154	318	35243.84	9.12	52.26	-0.636	"
Zm00001d037167	340	36604.44	5.16	69.24	-0.386	"
Zm00001d048796	219	24912.51	4.80	44.57	-0.769	"
Zm00001d039623	350	38476.97	5.76	90.54	0.057	"
Zm00001d018328	401	44022.06	7.23	83.24	-0.155	"
Zm00001d009571	512	56158.7	5.59	75.57	-0.319	"

The number of amino acids, molecular weight length, theoretical isoelectric point value, fat coefficient value and hydrophobic average coefficient value of each protein in the ZmBBI gene family are not the same in different ZmBBI gene families (Table 2). Among BBI proteins, Zm00001d035989 is the largest protein having 647 amino acids; while Zm00001d021225 being the smallest protein having 175 amino acids. Zm00001d039623 is with the highest fat coefficient of value 90.54 while Zm00001d048796 is with the smallest fat coefficient of value 44.57. The average hydrophobicity coefficient of ZmBBI family proteins in maize is negative, except Zm00001d039623, which indicated that most of them are hydrophilic proteins. There are 5 basic amino acids and 12 acid amino acids in 17 ZmBBI genes. According to the fat coefficient of ZmBBI family gene, the stability of ZmBBI family protein is strong.

Domain families on selected sequences





Fig. 2. Conserved domain of ZmBBI amino acid sequence in maize. Analysis and predicting the conserved domains of 17 ZmBBI proteins using online software Conserved Domain Database in NCBI database (<https://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi>).

Using the transcriptome sequencing data of maize tissues at different developmental stages published by Stelplflug *et al.* (2016), the expression of 17 ZmBBI family genes was obtained. An expression Heatmap was prepared at different stages of maize development and in different tissues

with the expression value of each gene ZmBBI family genes (Fig. 4). The expression patterns of ZmBBI family genes in different maize tissues and development stages are different. Zm00001d052154 and Zm00001d037167 showed no significant expression at any stages of the analysis, while Zm00001d048796, Zm00001d021314, Zm00001d039769, Zm00001d08649 and Zm00001d018328 showed moderate expression at all stages and tissues, Zm00001d04831 was highly expressed at most stages and tissues; Zm00001d03962 was highly expressed at mature pollen stage, but not expressed in other stages and tissues; Zm00001d021225 did not express in mature pollen and endosperm crown, but expressed highly in silk and cortex, and moderately expressed in other stages and tissues; Zm00001d04831 highly expressed in symmetrical division area, vegetative meristem and surrounding tissues, growth zone, internode sections 6 - 7 and 7-8, and moderately expressed in other stages and tissues. Zm00001d04831 was highly expressed in the whole process of maize growth and development and in various tissues, which shows that it is very important in plant growth.

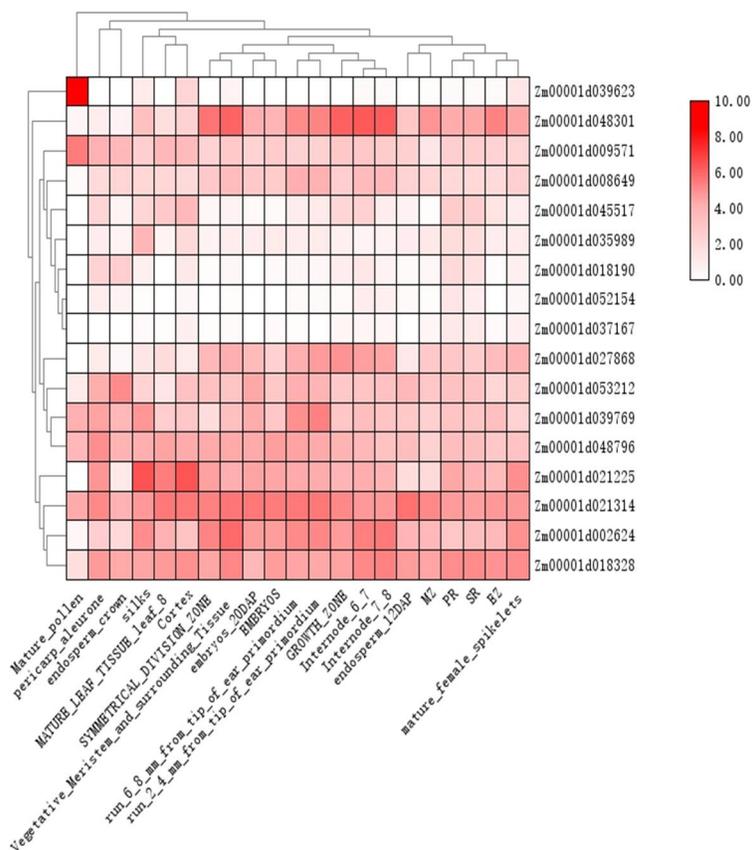


Fig. 3. Expression map of maize ZmBBI gene in different tissues of maize. Relative expression levels of ZmBBIs in different tissues of maize. The red represents expression at high level, and the white represents expression at low level.

Using the published data of salt and hypertonic stress in maize (Yu *et al.* 2016, Yu *et al.* 2018), the expression of 17 ZmBBI genes in leaves and roots under salt stress and hypertonic stress was also studied and drew a Heatmap that can be intuitively analyzed (Fig. 4). It can be seen

that the expression of ZmBBI family gene is not the same under different salt stress and hypertonic stress. GRMZm2g071602 (Zm00001d045517) was highly expressed under normal, hyperosmotic and salt stress conditions, GRMZM2g004799 (Zm00001d039623) was not expressed under all conditions, GRMZM2g085948 (Zm00001d021314) was highly expressed under most conditions, GRMZm2g099278 (Zm00001d035989), GRMZM2g473016 (Zm00001d052154) and GRMZM2g142816 (Zm00001d037167) were more expressed lower (Figs 3 and 4). The results showed that GRMZM2g165044 (Zm00001d039769) was expressed inefficiently under normal leaf conditions, but highly under salt stress and high osmotic stress, so it was speculated that abiotic stress had an important effect on the growth and development of maize.

All E3s have the ability to connect specific E2 to the target protein (Wang *et al.* 2020). The target receptor protein can be ubiquitinated by the ligase of ubiquitin chain, and the single ubiquitin can combine with other target eggs with ubiquitin binding domain white matter to change its position (Zhang *et al.* 2019, Dong *et al.* 2020)

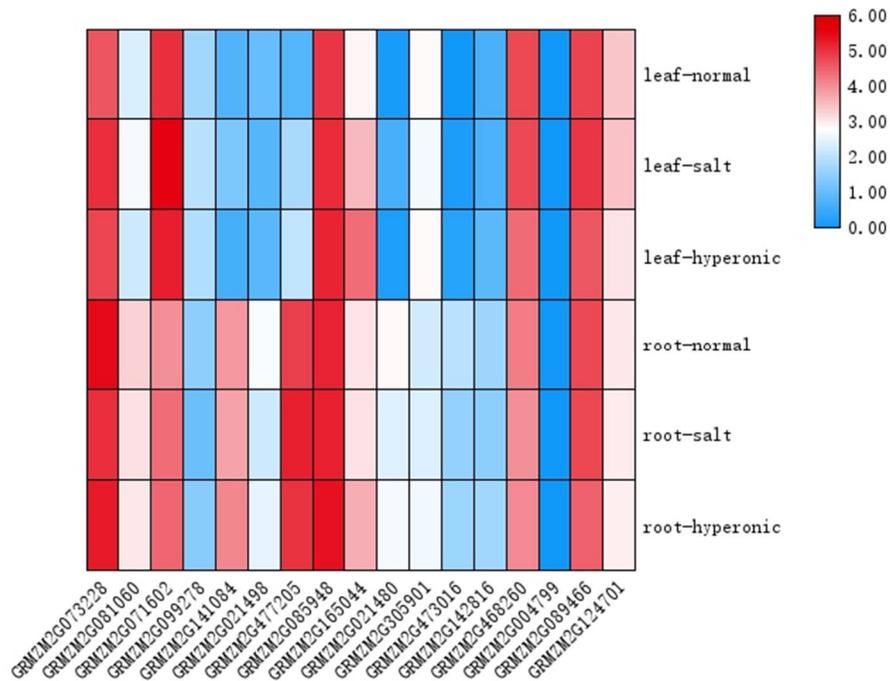


Fig. 4. Expression of ZmBBI gene in maize under salt and hypertonic stress. Relative expression levels of ZmBBIs in maize leaf and root under normal condition as well as salt- and hyperosmotic stress. The red represents expression at high level, and the blue represents expression at low level.

The results showed (Fig. 4) that GRMZM2g165044 (Zm00001d039769) was inefficient under normal conditions, but highly efficient under salt stress and high osmotic stress; GRMZM2g071602 (Zm00001d045517) was better under salt stress and high osmotic stress than under normal root conditions, so it was estimated that GRMZM2g165044 (Zm00001d039769) was more useful for the growth and development of maize under worse conditions.

In subcellular localization, it was found that all the ZmBBI genes are located in the nuclei (Table 2), but its function and mechanism of action are not clear yet, which needs to be studied further. The expression of protein was closely related to gene. Zm00001d021314 was highly expressed in the whole process of maize growth and development and in various tissues and organs, which indicates that this gene is very important in the whole plant life. Zm00001d039623 and Zm00001d009571 are well expressed in pollen, which indicates that these two genes may participate in the development of some plant organs; but Zm00001d037167 is not well expressed in many periods. Furthermore, under normal condition ZmBBI genes expression level maintain high levels, while the expression levels of ZmBBI genes are not changed under salt and hypertonic stress, which indicate that ZmBBI genes are very important under no matter in any condition. For example, Zm00001d039769 is not expressed well in normal leaves, but in any expressed in leaves under severe conditions, indicating that it plays an important role to cope with the severe conditions (salt and hypertonic stress) (Zhang *et al.* 2020)

Properties of the ZmBBI gene family in maize was depicted by bioinformatics means and results of such study may provide some basic data for further utilization of ZmBBI gene in maize breeding.

Acknowledgements

This work was supported by Science and Technology Bureau Outstanding Young Talent Development Project of Jilin City (#20200104064).

References

- Artimo P, Jonnalagedda M, Arnold K, Baratin D, Csardi G, de Castro E, Duvaud S, Flegel V, Fortier A, Gasteiger E, Grosdidier A, Hernandez C, Ioannidis V, Kuznetsov D, Liechi R, Moretti S, Mostaguir K, Redaschi N, Rossier G, Xenarios I, and Stockinger H 2012. ExPASy: SIB bioinformatics resource portal. *Nucleic Acids Res.* **40**(W1): W597-W603.
- Chou KC and Shen HB 2010. Plant-mPLoc: A top-down strategy to augment the power for predicting plant protein subcellular localization. *PLoS One.* **5**(6): e11335.
- Dong ZR, Zhou W, Sun D, Yan YC, Yang CC, Yang YF, Li HC, Zhi XT and Li T 2020. Role of the E3 ubiquitin ligase TRIM4 in predicting the PrognZmis of hepatocellular carcinoma. *J. Cancer* **11**: 247-252
- Jeanmougin F, Thompson JD, Gouy M, Higgins DG, and Gibson TJ 1998. Multiple sequence alignment with Clustal X. *Trends Biochem. Sci.* **23**: 403-405.
- Jiang HY 2010. Proteomic studies of proteasomal inhibitor treated undifferentiation and differentiation SH-SY5Y cells. Jin Lin University, China.
- Jiang L, Yu X, Chen D, Feng H and Li J 2020a. Identification, phylogenetic evolution and expression analysis of GATA transcription factor family in maize (*Zea mays*). *Intl. J. Agric. Biol.* **23**: 637-643.
- Jiang L, Yu X, Chen D, Wang H, Cao D, Zhang J and Li J 2020b. Expression analysis of trihelix gene family in maize (*Zea mays*) using bioinformatics. *Intl. J. Agric. Biol.* **23**: 863-868.
- Jin JP, Tian F, Yang DC, Meng YQ, Kong L, Luo JC and Gao G 2017. PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. *Nucleic Acids Research* **45**(D1): D1040-D1045.
- Li W 2010. A rice E3 ubiquitin ligase OsBBII contributes to broad-spectrum resistance against *Magnaporthe oryzae*. Zhe Jiang University, China.
- Liu C, XX Li, YL Su and YF Guo 2015. Genome-wide identification, phylogenetic analysis and expression profiling of the SBP transcription factor family in *Nicotiana tobacum*. *Chin. Tobacco Sci.* **36**:1-11.
- Lu S, Wang J, Chitsaz F, Derbyshire MK, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Marchler GH, Song JS, Thanki N, Yamashita RA, Yang M, Zhang D, Zheng C, Lanczycki CJ and Marchler-Bauer A. 2020. CDD/SPARCLE: the conserved domain database in 2020. *Nucleic Acids Res.* **48**(D1): D265-D268.

- Song M, Wang C, Wang H, Zhang T, Li JQ, Robert B, Lotfi C, Sun YH and Cui XJ 2020. Targeting ubiquitin protein ligase E3 component N-recognin 5 in cancer cells induces a CD8⁺ T cell mediated immune response. *Onco* **9**: 1245-1342.
- Stelpflug SC, Sekhon RS and Vaillancourt B 2016. An expanded maize gene expression atlas based on RNA sequencing and its use to explore root development. *Plant Genome* **9**: 85-94.
- Wang YP, Ren W, Wang RJ, Shao KZ, Gao HJ and Zhang JL 2020. Research advances on functions of SUMO E3 Ligase in plant abiotic stress adaptation. *Biol. Bull.* **36**:169-177.
- Xie T, Wang S, Huang L, Wang X, Kang LP and Guo LP 2014. Transcriptome-based bioinformatics analysis of *Arnebia euchroma* ERF transcription factor family. *Chin. J. Trad. Chin. Med.* **39**: 4732-4739.
- Yu XM, Jiang LL, R Wu, XC Meng, A Zhang, N Li, Q Xia, X Qi, JS Pang, ZY Xu and B Liu 2016. The core subunit of a chromatin-remodeling complex, ZmCHB101, plays essential roles in maize growth and development. *Sci. Rep.* **6**: 38504.
- Yu XM, Meng XC, Liu YT, Li N, Zhang A, Wang TJ, Jiang LL, Pang JS, Zhao XX, Qi X, Zhang MS, Wang SC and Liu B 2018. The chromatin remodeler ZmCHB101 impacts expression of osmotic stress-responsive genes in maize. *Plant Mol. Biol.* **97**: 451-465.
- Zhang YW, Zhang J, Zhu MJ, Zhang HL and Wang QW 2019. Homologous cloning of soybean E3 ubiquitin ligase gene GmAIRP1 and its functional Identification in tobacco. *J. Plant Gen. Res.* **20**: 1011-1019.
- Zhang HN, Liang LG, Hu D, Zhang YY, Zhang YJ, Shao HB, Zhao LN, Yang RP and Guo XL 2020. Functional characterization of maize heat shock transcription factor gene ZmHsf01 in thermotolerance. *Peer J.* **8**: 269-277.
- Zhao JR, Wang S, Li M, Lv HY, Wang DW, Ge YQ, Wei X and Yang WC 2018. Current status and perspective of maize breeding. *J. Plant Genet. Res.* **19**: 435-446.
- Zhao JL, Yao WJ, Wang SJ, Jiang YB and Zhou BR 2015. AP2/ERF gene family in *Populus trichocarpa* by bioinformatics. *J. Northeast For. Univ.* **43**: 21-29.
- Zheng H 2013. Functional analysis of zinc finger protein gene OsLOL2, OsC2HC-1 response to salinity stress in rice. Northeast Forestry University, China.
- Zhu S, Zeng L, Wu PZ, Chen YP, Jiang HW, Wu GJ and Li MR 2012. Cloning and sequence analysis of the cDNA of plastid-located glycerol-3-phosphate acyltransferase (JcGPAT2) gene from *Jatropha curcas* L. *J. Guangdong Agr. Sci.* **39**: 1-5.

(Manuscript received on 18 August, 2020; revised on 28 October, 2020)